

羊宇培

(+86)15221421028 | yupei_yang@sjtu.edu.cn | [yupei-yang.github.io](https://github.com/yupei-yang)

个人简介

上海交通大学计算机学院博士在读。近年来研究聚焦于**强化学习后训练 (LLM/RL Post-Training)** 与**大模型检索增强生成 (RAG)**，通过因果奖励建模缓解 RL 训练中的奖励欺骗与大模型“幻觉”问题，并探索其在信息抽取及多模态图文混排中的应用。相关工作的理论基础源于因果结构与强化学习的融合，旨在从因果视角提升强化学习系统的探索效率、泛化能力与决策可靠性。成果发表于 ICLR、IJCAI 等国际顶级会议，并在多个实际系统中完成验证。

教育经历

- 上海交通大学，计算机学院，认知机器和健康计算研究中心** 2022 – 至今
直博，导师：涂任奎（上海交通大学），合作导师：黄碧薇（加州大学圣地亚哥分校） 上海，中国
- 东南大学，吴健雄学院（荣誉学院），人工智能** 2018 – 2022
本科，学术导师：杨绍富（东南大学） 南京，中国

论文发表

- UniRank: End-to-End Domain-Specific Reranking of Hybrid Text-Image Candidates** [论文] [代码]
Yupei Yang, Lin Yang, Wanxi Deng, Lin Qu, Shikui Tu, Lei Xu
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2026, under review.
- Factored Causal Representation Learning for Robust Reward Modeling in RLHF** [论文] [代码]
Yupei Yang, Lin Yang, Wanxi Deng, Lin Qu, Fan Feng, Biwei Huang, Shikui Tu, Lei Xu
International Conference on Machine Learning (ICML), 2026, under review.
- Hallucination-Resistant Relation Extraction via Dependency-Aware Sentence Simplification and Two-tiered Hierarchical Refinement** [论文] [代码]
Yupei Yang, Fan Feng, Lin Yang, Wanxi Deng, Lin Qu, Biwei Huang, Shikui Tu, Lei Xu
Transactions of the Association for Computational Linguistics (TACL), 2025, under review.
- Towards Generalizable Reinforcement Learning via Causality-Guided Self-Adaptive Representations** [论文] [代码]
Yupei Yang, Biwei Huang, Fan Feng, Xinyue Wang, Shikui Tu, Lei Xu
International Conference on Learning Representations (ICLR), 2025.
- Boosting Efficiency in Task-Agnostic Exploration through Causal Knowledge** [论文] [代码]
Yupei Yang, Biwei Huang, Shikui Tu, Lei Xu
International Joint Conference on Artificial Intelligence (IJCAI), **Oral**, 2024.

实习经历

- 阿里巴巴** 2025.06 – 至今
学术合作实习生 · 大模型应用框架研发工程师 杭州
 - 围绕 **RL 后训练与多模态 RAG 系统优化** 开展研究，提升大模型在复杂垂域场景下的对齐能力与可靠性。
 - 探索因果奖励建模在 RLHF 中的应用，**缓解虚假相关与奖励欺骗问题**，提高模型对人类偏好的稳健刻画能力。
 - 研究多模态重排序中的“**模态鸿沟**”问题，提升图文混合候选场景下的检索精度与排序一致性。
 - 支持**学术搜索与企业私域知识问答**等真实业务场景。

项目经历

- 融合向量与知识图谱增强的多模态 RAG 检索生成系统项目 (阿里巴巴 × 上交)** 2025.06 – 至今
LLM, VLM, Causal Reward Modeling, RLHF, RAG
 - 项目背景:** 面向企业私域知识问答与学术搜索场景, 构建融合向量检索与知识图谱增强的混合式多模态 RAG 框架, 解决复杂跨模态检索与大模型幻觉问题。
 - 因果奖励建模:** 提出因果分解式奖励建模框架 CausalRM, 显式解耦因果与非因果表征因素, 提升 RLHF 训练稳定性与鲁棒性, 缓解奖励欺骗问题, 为后续信息抽取与检索重排模块提供可靠对齐信号, 相关成果投稿至 **ICML 2026**。
 - 多模态重排序:** 设计端到端多模态重排序模型 UniRank, 支持跨模态统一相关性评分, 并通过 SFT + RL 两阶段对齐实现领域自适应优化, 提升科研检索与专利搜索场景排序效果, 相关成果投稿至 **KDD 2026**。
 - 系统落地:** 完成多模态重排序模块的工程化部署与上线验证, 支持学术搜索与企业内部知识问答场景的真实业务应用。
- 面向金融智能问答的图数据库与 GraphRAG 攻关项目 (中国工商银行 × 华为云 × 上交)** 2024.09 – 2025.02
LLM-based Information Extraction, GraphRAG
 - 项目背景:** 面向金融监管与政策类长文档场景, 构建基于大模型的信息抽取与 GraphRAG 系统, 为金融智能问答提供结构化知识支撑。
 - 知识抽取与精细化构图:** 负责设计实体类型体系与抽取提示模板, 结合政策文档的章节条款结构进行结构感知建模, 开发循环抽取与上下文增强机制, 提高实体与关系覆盖率。
 - 图谱质量优化与系统迭代:** 构建实体合并与后处理流程, 引入相似实体匹配与多关系共存的图结构建模机制, 解决实体冗余、关系缺失与多文档同名冲突问题, 并针对漏抽、误抽与语义偏移问题进行多轮优化, 显著提升图谱完整性。
 - 研究延展:** 针对大模型抽取过程中的幻觉与误判问题, 进一步开展鲁棒抽取与幻觉抑制研究, 相关成果投稿至 **TACL**。
- 复杂多源环境下的因果强化学习推理与策略泛化项目 (空军研究院 × 上交)** 2023.11 – 2025.05
Causal RL, Causal Exploration, World Model, Domain Adaptation
 - 项目背景:** 面向复杂动态环境下的多源信息推理与决策规划任务, 构建基于因果结构建模的强化学习框架, 支持从单模态到多模态场景的策略泛化与自适应迁移, 获空军“数智”创客大赛三等奖。
 - 因果探索机制:** 提出基于因果图约束的任务无关探索方法, 通过选择因果父节点与剥离伪相关变量, 显著提升强化学习的样本效率与探索稳定性。相关成果发表于 **IJCAI 2024**。
 - 因果自适应表征与迁移:** 构建因果因子化世界模型与自适应策略表征框架, 实现跨场景策略快速迁移与鲁棒泛化。相关成果发表于 **ICLR 2025**。

获奖情况

- 优秀学生干部 (前 5%)** 2025.10
上海交通大学
- “榜样的力量” 优秀学生 (前 1%)** 2025.03
上海交通大学
- 优秀毕业生 (前 1%)** 2022.06
东南大学
- 校长奖学金 (前 5%)** 2019.10
东南大学

服务经历

- 中国共产党上海交通大学第十二次代表大会代表** 2025.06 – 至今
上海交通大学
- 博士生计算机第二党支部书记** 2023.09 – 2025.06
上海交通大学 · 电子信息与电气工程学院
- 健雄书院执行副院长 (第一届)** 2020.09 – 2021.06
东南大学 · 吴健雄学院

技能介绍

- 编程语言与工具:** Python, MATLAB, C++; PyTorch, TensorFlow; OpenRLHF, Swift.